

# 文単位の Nbest 候補制約に基づく文書翻訳

駒田 啓伍<sup>1</sup> 森下 睦<sup>2</sup> 鈴木 潤<sup>1</sup><sup>1</sup> 東北大学 <sup>2</sup> NTT コミュニケーション科学基礎研究所  
keigo.komada.r6@dc.tohoku.ac.jp

## 概要

現在の機械翻訳モデルは主に1文から1文への翻訳を対象としており、単文に対しては高い精度を達成できるものの、文書全体としての表現の一貫性などに関しては課題が残っている。また従来の文書翻訳手法は、文書から文書への長い系列変換を行うことによる翻訳精度低下など実用上の問題が残っていた。本研究では、この問題を軽減するために、単文翻訳モデルと文書翻訳モデルの両方を活用する手法を提案する。具体的には、単文翻訳モデルが出力した翻訳候補をもとに、文書翻訳モデルによって適切な翻訳文の組合せを求め文書翻訳を実現する。実験により、従来の文書翻訳手法と比較して提案手法は高い文書翻訳精度が得られることが示された。

## 1 はじめに

機械翻訳技術の進展により、機械翻訳による翻訳品質は実用レベルに達し、既に商用システムとして一般にも広く使われている。これにより、機械翻訳の研究が完結したかという点、必ずしもそうではない。一文単位の翻訳（以下「**単文翻訳**」と略記）、特にデータが集めやすいドメインの文章の翻訳に関しては十分に実用レベルと言えるが、例えば文章全体を翻訳する際などに表現やスタイルが一貫した翻訳結果を得ることは、現状の機械翻訳技術でも容易ではない。実際に、機械翻訳の研究コミュニティにおいても、文書単位で一貫性のある翻訳（以下「**文書翻訳**」と略記）の実現は、喫緊に取り組むべき課題として認識されている。

一貫性のある文書翻訳結果を得たい場合、安直には文書全体を一文とみなして処理すれば、原理的には単文翻訳モデルでも翻訳することはできる。しかしこういった方法は一般的には現実的ではない。それは、現在広く用いられている Transformer モデルによるニューラル翻訳は、本質的に文長が長くなるとう性能が急激に悪くなる性質を持つからである [1]。

また、別観点の問題として、高い翻訳品質を達成するには多くのデータを必要とすることが一般的である。しかし、文書単位の対訳データは、文単位の対訳データに比べてそれほど多くない。これらの現状から、現在の文単位のニューラル翻訳をそのまま文書翻訳に転用しても、良い結果が得られない。

これらの背景から、本稿では文書翻訳の性能向上を目標に新たな方法を提案する。具体的には、従来用いられてきた Transformer ベースの単文翻訳モデルに対して、少量の文書対訳データを用いて追加学習<sup>1)</sup>を行うことで、データ量の問題に対応する。また、単文翻訳モデルが生成する複数の翻訳候補を文書翻訳時の制約と考え、文単位の制約付き生成をすることで長文を生成する際の翻訳品質の劣化を防ぎ文書単位の翻訳を獲得する。実験では、chat 及び news 文書に対して実験を行い、単文翻訳で各文を独立に翻訳した結果を並べる形式で文書を翻訳するよりも、BLEU スコアの意味で高い翻訳精度を達成したことを報告する。

## 2 文書翻訳

機械翻訳システムの翻訳結果を文書全体で評価した時、整合性や一貫性の観点で人間に及ばない場合が多いことが知られている [2]。これは、一般的なニューラル機械翻訳モデルが1文から1文への翻訳を対象としているため、文書全体で見た際に表現やスタイルの一貫性が欠如しがちなことが主要要因として挙げられる。

この問題を解決するために、翻訳時に前後の文（文脈）を考慮する文書翻訳の手法がいくつか提案されている。例えば、Tiedemann&Scherrer ら [3] は、入力文と出力文に1つ前の文を結合し、2文ずつモデルに入力することで文脈を考慮する手法を提案した。また、Junczys-Dowmunt [4] はこれを拡張し、文書全体を入出力とするモデルを提案した。これらの手法は、シンプルかつ効果的であることが検証され

1) 本稿では fine-tuning を「追加学習」と表記する

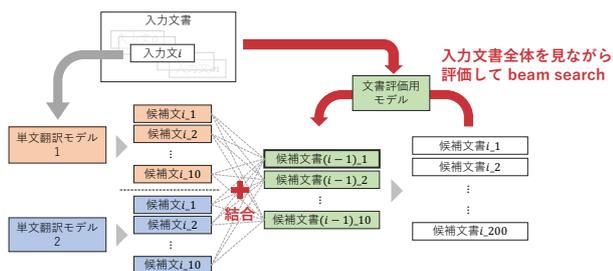


図1 文書翻訳のアルゴリズムの概略図 ( $N = 10$  の場合)

ているものの、現在の Transformer モデルは文長が長くなるにつれ計算量が大幅に増加し、翻訳精度が低下する欠点 [1] が指摘されており、実用上の問題が残っている。

本研究の提案法は、従来法の課題である翻訳時の計算量を抑えつつ文書全体を考慮した翻訳を実現する方法となる。次節以降では、提案手法の詳細を説明する。

### 3 提案法

本研究では、通常の 1 文から 1 文へ翻訳を行うモデル (単文翻訳モデル) と、文書全体を入力し文書全体を出力するモデル (文書翻訳モデル) の両方を活用することを考える。具体的には、単一ドメインの対訳コーパスで追加学習した単文翻訳モデルで文書中の各文を独立に翻訳し翻訳候補を得た上で、それらを文書単位の対訳コーパスで追加学習した文書翻訳モデルの生成制約とする。単文翻訳モデルを単一ドメインに特化させるように追加学習したことで、ドメイン不明の入力文に対してもその文のドメインと一致する単文翻訳モデルが存在していれば、よい翻訳を出力できるようになることを期待している。これにより、文書翻訳モデルは新たに文を生成するのではなく、単文翻訳モデルで得られた翻訳文をもとに文書全体として適切な文の組合せを求めることとなる。

以下に本手法の詳細な生成手順を述べる。図 1 のように単一ドメインの対訳コーパスで追加学習した単文翻訳モデルを複数個、文書翻訳モデルを 1 つ用意した。以下において  $i$  は処理対象の単文が文書内において  $i$  番目であることを表す。

1. **候補文を生成** 入力文書内の  $i$  番目の文を翻訳して、各単文翻訳モデルから  $N$  個ずつ候補文を出力させる。
2. **候補文書を生成** 前の  $(i-1)$  番目ループで残った

表 1 追加学習用及び評価用データの統計量。

データ	En-Ja		Ja-En		
	文書数	文数	文書数	文数	
追加学習用	news	140	1,998	140	1,998
	chat	143	2,096	199	2,198
評価用	news	63	1,000	80	993
	chat	57	644	51	542

候補文書の末尾に候補文を結合する。この操作を全ての候補文書と候補文の組合せに対して行って新たな候補文書とする。 $i=0$  の時は、前のループで残った候補文書は空集合とする。

3. **候補文書の評価** 入力文書の全体を参照しながら、文書翻訳モデルで候補文書の評価する
4. **beam search** 評価して得られた尤度をもとに、候補文書から  $N$  個だけ選択して次のループに残す。1 に戻ってループする。

**ループ終了** 全ての文に対して操作が終了したら、 $N$  個の候補文書のうち最も尤度が高かったものを提案法の出力とする。

beam search によって Nbest 候補制約を与えたのは、制約しない場合はと生成した文書数が、文数に対して指数関数的に増加して、生成した文書の評価をするのに極めて長い計算時間を必要するからである。

### 4 実験

**データ** 追加学習用および評価用データとして、文書単位の情報をもつ chat および news ドメインの対訳コーパスを用意した、表 1 にそれらの文書数および文数を示す。chat ドメインの対訳コーパスとして BPersona-chat [5] を用いた。同様に、news ドメインの対訳コーパスとして WMT-2020 ニュース翻訳タスクの開発/評価セット (newsdev2020, newstest2020) [6] を使用した。BPersona-chat は文書の単位でランダムに一部を評価用として切り出し、残りを追加学習用とした。また、news ドメインは newsdev2020 を追加学習用、newstest2020 を評価用とした。

**モデル** 単文翻訳の事前学習済み翻訳モデルとして、WMT-2022 で好成績を得た NT5 チームが学習した Transformer [7] ベースの翻訳モデル [8] を用いた。以下、この事前学習済みモデルをベースモデルと呼ぶ。次に、このベースモデルから独立に 4 つの翻訳モデルを追加学習により獲得した。

**表 2** 文書翻訳モデル DocMT(news&chat) と単文翻訳モデル SentMT(news&chat) の文書翻訳結果を sacrebleu で評価した際の BLEU スコア

モデル	En-Ja		Ja-En	
	news	chat	news	chat
DocMT(news&chat) : 提案法	<b>28.2</b>	<b>19.0</b>	<b>26.5</b>	30.3
SentMT(news&chat) : ベースライン	27.7	17.9	26.2	<b>30.8</b>

**表 3** 文書翻訳モデル DocMT(news&chat) で単文翻訳と同じ翻訳をした場合との BLEU スコア

モデル: DocMT(news&chat)	En-Ja		Ja-En	
	news	chat	news	chat
単文翻訳と同等の翻訳	26.9	14.7	25.9	28.8
DocMT(news&chat) からの差分	-1.3	-4.3	-0.6	-1.5
SentMT(news&chat) からの差分	-0.8	-3.2	-0.3	-2.0

1. **DocMT(news&chat) : news 及び chat** ドメイン両方の文書対訳コーパスを使って追加学習をした**文書翻訳**モデル
2. **SentMT(news&chat) : news 及び chat** ドメイン両方の文書対訳コーパスから文書情報を取り除き文単位の追加学習をした**単文翻訳**モデル
3. **SentMT(news) : news** ドメインの文書対訳コーパスから文書情報を取り除き文単位の追加学習をした**単文翻訳**モデル
4. **SentMT(chat) : chat** ドメインの文書対訳コーパスから文書情報を取り除き文単位の追加学習をした**単文翻訳**モデル

文書翻訳モデルは Junczys-Dowmunt [4] の手法を参考に、文書中の各文を結合し学習した。提案法の説明で述べたように、SentMT(chat) と SentMT(news) は、候補の生成のために利用する。また、DocMT が主たる文書翻訳モデルである。前記 SentMT(chat) と SentMT(news) モデルが生成した候補を制約として文書翻訳をおこなう。SentMT(news&chat) は、提案法 DocMT(news&chat) と比較して文書単位で学習したか文単位で学習したかの違いだけのモデルになる。なお翻訳モデルの学習には fairseq ツールキット [9] を使用した。また、文書翻訳を実行するために fairseq ツールキットを活用しつつ制約付き翻訳用のコードを独自に作成した。

**評価指標** 本実験の評価は、WMT-2020 の標準評価ツールである sacrebleu [10] を用いて BLEU スコア [11] にて実施した。

**表 4** 提案法の候補文を生成するために用いた、ドメインごとに追加学習した単文翻訳モデルの BLEU スコア

モデル	En-Ja		Ja-En	
	news	chat	news	chat
SentMT(chat)	26.4	<b>18.7</b>	24.5	<b>30.1</b>
SentMT(news)	<b>27.7</b>	12.6	<b>26.4</b>	26.3

## 4.1 実験結果

表 2 に主たる実験結果を示す。Ja-En の chat 評価データに対する翻訳結果を除いて、提案法の BLEU スコアの方が高いという結果になった。この結果から、提案法である文書翻訳は、単文翻訳モデルが文書内の文を一文ずつ翻訳して結合する方法（ベースライン）により文書を翻訳するよりも、概ね良好な結果が得られることがわかった。

表 3 に、DocMT(news&chat) に対して単文翻訳から得た候補を用いずに、通常の単文翻訳と同様に翻訳した際の結果を示す。また、表 2 の DocMT(news&chat) および SentMT(news&chat) からの BLEU スコアの差分を示す。差分が全てマイナスの値になっているのは、DocMT(news&chat) で単文翻訳と同じ翻訳をすると、ベースラインである SentMT(news&chat) よりもさらにスコアが悪くなることを示している。この結果から、提案法で導入した単文翻訳の候補を制約として文書翻訳に利用することの有効性が示された。使うモデルが完全に同じでも、制約を入れるか入れないかでこのような大きな違いが見られた理由は、文書翻訳の場合は、翻訳文の文長が単文翻訳に比べて長くなることが要因と考えられる。例えば、学習データに現れない文長の翻訳をするとき、極端に翻訳品質が悪くなることが知られている [12, 13]。また、モデル学習時と翻訳時の状況のずれいわゆる exposure bias に起因する要因もあると考えられる [14, 1]。

## 4.2 分析/考察

**単一のドメインの学習結果** 表 4 に、提案法である DocMT の候補を生成した単文翻訳 SentMT(chat) と SentMT(news) の 1 ベスト翻訳の性能を示す。chat 評価データを SentMT(chat) にて翻訳する場合、あるいは、news 評価データを SentMT(news) にて評価する場合のようにモデルと評価データのドメインが一致するときは高いスコアが得られた。反対にモデルと評価データのドメインが一致しない場合は

**表 5** 候補文のドメインの内訳と、単文翻訳モデルが候補文を生成した段階での候補文の順位 (0 位から) についての各種統計量 (chat ドメインの対訳コーパスにおける場合)

モデル: SentMT		En-Ja		Ja-En	
		chat	news	chat	news
内訳	個数	786	58	529	13
	混合率	0.931	0.069	0.976	0.024
	最頻値	0	2	0	0
候補文の順位	平均	3.113	3.552	2.694	2.846
	標準偏差	2.622	2.848	2.655	2.905

BLEU スコアが大幅に悪くなった。もう一点着目すべきこととして、表 3 の DocMT(news&chat) の結果は、評価データとモデルのドメインが一致している場合の結果よりもわずかに良い結果となっている。この結果から、DocMT(news&chat) は与えられた候補の中からより良い翻訳候補を選択できたことが示唆される。

**モデルの選択結果の統計量** 提案法が実際に出力した文書を構成した候補文に注目して、単文モデルが生成した候補文のドメインの内訳、モデルが候補文を生成した段階での候補文の尤度による順位に関するいくつかの統計量の 2 つについて、表 5 に chat ドメインの文書を提案法が翻訳した場合、表 6 に news ドメインの文書を提案法が翻訳した場合を示す<sup>2)</sup>。選択された候補文の内訳に注目すると、chat ドメインのデータを翻訳した場合は SentMT(news) 由来の候補文があまり混在しなかったのに対して、news ドメインのデータを翻訳した場合は SentMT(chat) 由来の候補文が 30% 以上も混在していた。選択した候補文の順位について、共通する傾向としてどちらも順位の高い (0 に近い側) 候補文を選択したことが分かった。平均は news ドメインの文書を翻訳した場合の方が低かった。また翻訳する文書と候補文のドメインが異なる場合は、選択された候補文の順位の標準偏差は大きくなる傾向が見られた。

**翻訳結果の一貫性の定性評価** 図 2 に提案法による実際の chat 文書翻訳結果を示す。この出力の中には SentMT(news) の候補文が含まれているが、その候補文でも chat ドメインのようによく似た印象の語彙が使われているのが分かった。一方で SentMT(news) に限らず、SentMT(chat) の候補文の一部であってもです・ます調の部分があり、正解

2) 選択された候補文の尤度による順位に関する分布や、文書内で何文目なのかという位置についての情報の詳細は付録の表 7, 9, 10, 11, 12 を参照

**表 6** 候補文のドメインの内訳と、単文翻訳モデルが候補文を生成した段階での候補文の順位 (0 位から) についての各種統計量 (news ドメインの対訳コーパスにおける場合)

モデル: SentMT		En-Ja		Ja-En	
		chat	news	chat	news
内訳	個数	332	648	439	554
	混合率	0.339	0.661	0.442	0.558
	最頻値	1	1	0	0
候補文の順位	平均	4.015	3.640	2.957	3.153
	標準偏差	2.927	2.883	2.686	2.682

文数	ドメイン	翻訳結果
0	chat	こんにちは、何かペットは飼ってますか?
1	news	こんにちは!はじめまして。はい、犬を 4 匹飼っています。
2	news	いいですね、私はどちらかという猫派ですが、23 匹飼っています。
3	news	兄弟や姉妹はいますか?兄弟がいないので犬を飼っています。
4	chat	はい、でも頻繁には会いません、私はボートで世界中を旅しています。
5	chat	私は旅行が大好きです!ヨーロッパとメキシコには行ったことがあります。あなたはどこに旅行しますか?
6	chat	世界中を旅していろんなヴィーガン料理を試しています。
7	news	メキシコにいるときは、ホットチリや辛い食べ物が大好きです!あなたは?
8	news	はい、魚に限らず、いろんな種類の食べ物が好きです。

**図 2** 提案法による Chat 文書の En-Ja 翻訳結果

データのような表現が他の SentMT(chat) の候補文に存在していたがそれを選択していなかった。また、誤誤も見られた。このように、提案法によってスコアが向上したが、まだ改善すべき点は残っていることが分かった。

## 5 おわりに

本研究では、単文翻訳モデルが出力した翻訳候補をもとに、文書翻訳モデルによって適切な翻訳文の組合せを求め、文書翻訳を実現する手法を提案した。結果として chat 及び news ドメインの対訳コーパスに対してこの提案法は、単文翻訳の結果を結合して文書を構成する手法よりも自動評価においてより高いスコアを得ることができた。しかし実際の出力を見ると、文書内での口調・語尾の使用の統一は完全になされているわけではなく、誤誤も確認された。

簡単に準備できて、それなりに品質の高いデータが今回用いた chat と news データしかなかったため 2 種類のデータのみで実験を行ったが、より広範な性能を評価するためには、評価データを増やす必要がある。今後、文書翻訳の性能を高めるために、これらの評価データおよび追加学習用のデータを構築していきたいと考えている。

## 謝辞

本研究は JST ムーンショット型研究開発事業 JPMJMS2011 (fundamental research) の助成を受けて実施されたものである。

## 参考文献

- [1] Marc'Aurelio Ranzato, Sumit Chopra, Michael Auli, and Wojciech Zaremba. Sequence level training with recurrent neural networks. In **4th International Conference on Learning Representations, ICLR 2016**.
- [2] Samuel Läubli, Rico Sennrich, and Martin Volk. Has machine translation achieved human parity? a case for document-level evaluation. In **Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing**, pp. 4791–4796, Brussels, Belgium, October–November 2018. Association for Computational Linguistics.
- [3] Jörg Tiedemann and Yves Scherrer. Neural machine translation with extended context. In **Proceedings of the Third Workshop on Discourse in Machine Translation**, pp. 82–92, Copenhagen, Denmark, September 2017. Association for Computational Linguistics.
- [4] Marcin Junczys-Dowmunt. Microsoft translator at WMT 2019: Towards large-scale document-level neural machine translation. In **Proceedings of the Fourth Conference on Machine Translation (Volume 2: Shared Task Papers, Day 1)**, pp. 225–233, Florence, Italy, August 2019. Association for Computational Linguistics.
- [5] Yunmeng Li, Jun Suzuki, Makoto Morishita, Kaori Abe, Ryoko Tokuhisa, Ana Brassard, and Kentaro Inui. Chat translation error detection for assisting cross-lingual communications. In **Proceedings of the 3rd Workshop on Evaluation and Comparison of NLP Systems**, pp. 88–95, Online, November 2022. Association for Computational Linguistics.
- [6] Loïc Barrault, Magdalena Biesialska, Ondřej Bojar, Marta R. Costa-jussà, Christian Federmann, Yvette Graham, Roman Grundkiewicz, Barry Haddow, Matthias Huck, Eric Joanis, Tom Kocmi, Philipp Koehn, Chi-kiu Lo, Nikola Ljubešić, Christof Monz, Makoto Morishita, Masaaki Nagata, Toshiaki Nakazawa, Santanu Pal, Matt Post, and Marcos Zampieri. Findings of the 2020 conference on machine translation (WMT20). In **Proceedings of the Fifth Conference on Machine Translation**, pp. 1–55, Online, November 2020. Association for Computational Linguistics.
- [7] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. In **Advances in Neural Information Processing Systems**, Vol. 30, 2017.
- [8] Makoto Morishita, Keito Kudo, Yui Oka, Katsuki Chousa, Shun Kiyono, Sho Takase, and Jun Suzuki. Nt5 at wmt 2022 general translation task. In **Proceedings of the Seventh Conference on Machine Translation**, pp. 318–325, Abu Dhabi, December 2022. Association for Computational Linguistics.
- [9] Myle Ott, Sergey Edunov, Alexei Baevski, Angela Fan, Sam Gross, Nathan Ng, David Grangier, and Michael Auli. fairseq: A fast, extensible toolkit for sequence modeling. In **Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics (Demonstrations)**, pp. 48–53, Minneapolis, Minnesota, June 2019. Association for Computational Linguistics.
- [10] Matt Post. A call for clarity in reporting BLEU scores. In **Proceedings of the Third Conference on Machine Translation: Research Papers**, pp. 186–191, Brussels, Belgium, October 2018. Association for Computational Linguistics.
- [11] Kishore Papineni, Salim Roukos, Todd Ward, and Wei-Jing Zhu. Bleu: a method for automatic evaluation of machine translation. In **Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics**, pp. 311–318, Philadelphia, Pennsylvania, USA, July 2002. Association for Computational Linguistics.
- [12] Masato Neishi and Naoki Yoshinaga. On the relation between position information and sentence length in neural machine translation. In **Proceedings of the 23rd Conference on Computational Natural Language Learning (CoNLL)**, pp. 328–338, Hong Kong, China, November 2019. Association for Computational Linguistics.
- [13] Shun Kiyono, Sosuke Kobayashi, Jun Suzuki, and Kentaro Inui. SHAPE: Shifted absolute position embedding for transformers. In **Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing**, pp. 3309–3321, Online and Punta Cana, Dominican Republic, November 2021. Association for Computational Linguistics.
- [14] Wen Zhang, Yang Feng, Fandong Meng, Di You, and Qun Liu. Bridging the gap between training and inference for neural machine translation. In **Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics**, pp. 4334–4343, Florence, Italy, July 2019. Association for Computational Linguistics.

**表 7** SentMT(chat) が chat ドメインの対訳コーパスにおいて En-Ja 翻訳する際に出力した各順位の候補文が選択された回数と、文書内で何文目に選択されたかという位置に関する各種統計量

順位	選択された回数	最頻値	文書内での位置		
			平均	分散	標準偏差
0	150	0	6.35	21.35	4.62
1	126	1	6.59	20.40	4.52
2	109	10	7.37	19.90	4.46
3	90	5	6.86	16.75	4.09
4	83	7	7.31	19.08	4.37
5	68	11	7.41	19.18	4.38
6	54	8	7.59	18.28	4.28
7	41	8	7.15	17.93	4.23
8	35	4	6.29	17.92	4.23
9	30	4	7.83	15.47	3.93

**表 8** SentMT(chat) が chat ドメインの対訳コーパスにおいて Ja-En 翻訳する際に出力した各順位の候補文が選択された回数と、文書内で何文目に選択されたかという位置に関する各種統計量

順位	選択された回数	最頻値	文書内での位置		
			平均	分散	標準偏差
0	142	1	5.05	10.22	3.20
1	99	0	3.97	10.55	3.25
2	62	5	5.19	9.19	3.03
3	53	6	5.15	8.17	2.86
4	43	2	4.84	8.28	2.88
5	31	0	4.74	10.32	3.21
6	29	8	5.45	10.66	3.27
7	31	4	5.23	8.56	2.93
8	26	4	5.27	8.89	2.98
9	13	0	4.08	9.46	3.07

**表 9** SentMT(chat) が news ドメインの対訳コーパスにおいて En-Ja 翻訳する際に出力した各順位の候補文が選択された回数と、文書内で何文目に選択されたかという位置に関する各種統計量

順位	選択された回数	最頻値	文書内での位置		
			平均	分散	標準偏差
0	44	0	7.45	30.29	5.50
1	47	0	9.36	44.74	6.69
2	33	9	8.97	37.18	6.10
3	31	0	10.74	78.00	8.83
4	39	2	9.64	47.67	6.90
5	29	2	10.45	41.76	6.46
6	24	18	9.92	42.08	6.49
7	26	6	9.46	33.86	5.82
8	31	5	10.90	39.25	6.26
9	28	1	8.43	39.89	6.32

**表 10** SentMT(news) が news ドメインの対訳コーパスにおいて En-Ja 翻訳する際に出力した各順位の候補文が選択された回数と、文書内で何文目に選択されたかという位置に関する各種統計量

順位	選択された回数	最頻値	文書内での位置		
			平均	分散	標準偏差
0	101	6	9.12	47.45	6.89
1	104	3	8.14	37.99	6.16
2	76	9	9.04	32.96	5.74
3	71	9	9.28	46.62	6.83
4	52	7	8.44	43.48	6.59
5	56	4	7.77	45.54	6.75
6	51	4	9.12	54.57	7.39
7	45	6	8.40	38.42	6.20
8	47	3	6.62	28.62	5.35
9	45	1	9.13	52.74	7.26

**表 11** SentMT(chat) が news ドメインの対訳コーパスにおいて Ja-En 翻訳する際に出力した各順位の候補文が選択された回数と、文書内で何文目に選択されたかという位置に関する各種統計量

順位	選択された回数	最頻値	文書内での位置		
			平均	分散	標準偏差
0	98	3	6.90	30.85	5.55
1	73	5	8.52	42.39	6.51
2	62	3	8.03	30.26	5.50
3	46	5	7.30	16.52	4.06
4	42	4	8.02	34.79	5.90
5	35	2	7.63	42.18	6.49
6	19	1	10.53	43.09	6.56
7	22	10	8.59	42.06	6.49
8	26	6	8.85	30.98	5.57
9	16	5	8.13	38.73	6.22

**表 12** SentMT(news) が news ドメインの対訳コーパスにおいて Ja-En 翻訳する際に出力した各順位の候補文が選択された回数と、文書内で何文目に選択されたかという位置に関する各種統計量

順位	選択された回数	最頻値	文書内での位置		
			平均	分散	標準偏差
0	112	0	6.81	38.76	6.23
1	83	1	5.95	25.52	5.05
2	75	0	6.76	30.29	5.50
3	63	0	5.95	30.17	5.49
4	52	2	6.50	39.75	6.30
5	53	3	7.04	34.22	5.85
6	32	1	6.84	34.32	5.86
7	31	1	5.58	25.66	5.07
8	34	1	6.97	38.38	6.20
9	19	1	9.00	46.00	6.78